Data Science und Sicherheit

In «Data Science und Sicherheit» beschäftigen sich die Lernenden mit der digitalen Darstellung von Informationen, Datenverschlüsselung, Komprimierung, selbstkorrigierenden Kodierungen, Datenmanagement und maschinellem Lernen. Sie setzen sich nicht nur mit der Umsetzung dieser Konzepte auseinander, sondern auch mit dem geschichtlichen Kontext, in den diese Themenbereiche eingebettet sind.

Neu erschienen!

DIESE INHALTE STEHEN IHNEN ONLINE ZUR VERFÜGUNG:

- Lösungen zu den Aufgaben im Buch
- Übungsdateien (Excel) zu «Datenmanagement» und «Aus Daten lernen»
- Lernumgebungen zum Üben wichtiger Konzepte
- E-Tutorials
- Weitere Unterrichtsmaterialien

AUSZUG AUS «DATENMANAGEMENT»

Neue Konzepte und Begriffe

Eine Sammlung von Daten entspricht meistens einer grossen Menge von Tupel (Datensätzen). Die Aufgabe ein Tupel mit einer vorgegebenen Eigenschaft (z. B. den vorgegebenen Wert eines Attributs) zu finden, nennen wir den **Abruf des Tupels**.
Die Suche nach einem Datensatz mit bestimmtem Wert eines Attributs kann schneller bewältigt werden, wenn die Daten nach diesem Attribut sortiert sind. Wenn man Datensätze nach einem spezifischen Attribut sortiert, sprechen wir von **Indexierung** von Daten. Die resultierende Liste nennen wir **Index**. Die bekannte binäre Suche bietet hier die schnellste Strategie zum Finden spezifischer Datensätze.

5.7 Können Sie alle Personen finden, die jünger als 30 Jahre sind? Welche der Tabellen in Beispiel 5.1 benutzen Sie zu diesem Zweck?

5.8 Können Sie schätzen, was die folgenden SQL-Anfragen bewirken?

a
```
SELECT *
FROM people
ORDER BY Mitgliedsnummer
```

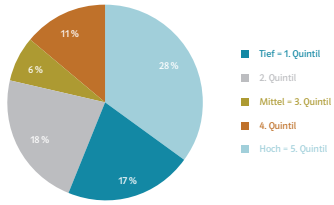
b
```
SELECT *
FROM people
ORDER BY Name ASC, Vorname DESC
```

AUSZUG AUS «AUS DATEN LERNEN»

Kontrollaufgaben zu: Informationen aus Datensammlungen visualisieren

6.9 In den folgenden Grafiken ist der Diagrammtyp nicht optimal gewählt. Geben Sie in jeder Grafik an, welcher Typ besser geeignet wäre, die Daten darzustellen.

c Die Abbildung unten zeigt den Anteil der Babys in Vietnam, die in ihren ersten sechs Monaten nur durch Muttermilch ernährt wurden, nach Einkommensklasse (fünf Einkommensklassen, die jeweils gleich viele Einwohner umfassen, dies nennt man Quintile).

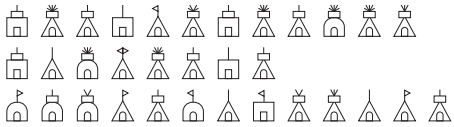


Quelle: Multiple Indicator Cluster Survey (MICS) 2006–2011

ZUSAMMENSPIEL VON AUFGABEN, LERNUMGEBUNG UND LÖSUNGEN – AUSZUG AUS «GEHEIMSCHRIFTEN UND DATENSICHERHEIT»

1. Eine Aufgabe lösen

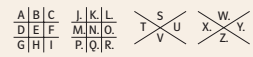
2.10 Der folgende Geheimtext basiert auf der Erzeugung von neuen Zeichen mit einer 3 × 9-Tabelle. Dechiffrieren Sie den folgenden Geheimtext, wenn bekannt ist, dass E am häufigsten, I am zweithäufigsten und die Buchstaben W und A genau dreimal vorkommen.



Geschichtlicher und gesellschaftlicher Kontext

Die Entwicklung der Kryptologie als Lehre der Geheimschriften ist stark verbunden mit dem Begriff der Sicherheit, d.h. in der Antike mit der Frage: «Wann ist eine Geheimschrift sicher?» Das Wort «sicher» soll bedeuten, dass kein Uneingeweihter es schafft, die Geheimtexte zu dechiffrieren.

Interessanterweise beruhte im Altertum die Sicherheit fast ausschliesslich auf der Geheimhaltung der Geheimschrift. Es wurden noch keine Methoden entwickelt, mit denen man aus den Geheimtexten selbst herausfinden könnte, wie die Geheimschrift funktioniert. Deswegen war das oberste Prinzip der Sicherheit, die verwendete Geheimschrift geheim zu halten. In der Praxis bedeutete das meistens, dass man die Beschreibung der Geheimschrift nicht schriftlich aufbewahren durfte. Deswegen waren alle Geheimschriften so gebaut, dass man sie leicht auswendig lernen konnte.



Die Geheimschrift der Freimaurer im Mittelalter ist das wohl bekannteste Beispiel dieser Art. Sie wurde von antiken Geheimschriften aus dem Gebiet des alten Palästina abgeleitet. Die Freimaurer zeichneten ihre Chiffrierungstabellen in Sand oder in Mehl während des Lernprozesses, um das aufgezeichnete Geheimnis nach dem Erlernen leicht verwischen, also unkenntlich machen zu können.

2. Die verfügbare Lösung konsultieren

2.10 **Lösung**: Zuerst bemerken wir, dass die Zeichen des Geheimtextalphabets 3 unterschiedliche Basisteile haben und neun unterschiedliche obere Teile. Das passt gut zu einer 3 x 9-Tabelle. Am häufigsten kommt  vor und somit ist es die Kodierung von E. Das zweithäufigste Zeichen ist  , das I kodiert. Somit ist  die Bezeichnung der ersten Zeile,  die Bezeichnung der fünften Spalte und die Bezeichnung der neunten Spalte.

	A	B	C	D	E	F	G	H	I
	J	K	L	M	N	O	P	Q	R
	S	T	U	V	W	X	Y	Z	

Wir könnten jetzt gut das System der Bezeichnungen der Spalten 5 bis 9 schätzen und somit einen grossen Schritt in der Kryptoanalyse machen. Oder wir setzen E und I ein und untersuchen den folgenden Lückentext:

– I E – – – – I E – I – – – – – I E – E – – – – E – – – – – – – E

Wenn wir als erstes Wort WIE schätzen (was gut zu dem angekündigten dreifachen Vorkommen von W passt und zusätzlich der obere Teil des Zeichens mit der Bezeichnung der fünften Spalte, in der W liegt, übereinstimmt), kodiert das dreimal vorkommende Symbol den Buchstaben A. Somit kennen wir jetzt die Bezeichnungen für alle drei Zeilen und zusätzlich für die erste Spalte.

	I								
	A	B	C	D	E	F	G	H	I
	J	K	L	M	N	O	P	Q	R
	S	T	U	V	W	X	Y	Z	

Jetzt kennen wir die Kodierungen von 8 Buchstaben A, J, S, E, N, W, I und R.

Die Dechiffrierung des entstandenen Lückentextes kann man jetzt leicht vollenden:

W I E S – – W I E R I – W A R – I E S E – N – – E – A – – – A – E

Die Chiffrierungstabelle sieht wie folgt aus:

	I								
	A	B	C	D	E	F	G	H	I
	J	K	L	M	N	O	P	Q	R
	S	T	U	V	W	X	Y	Z	

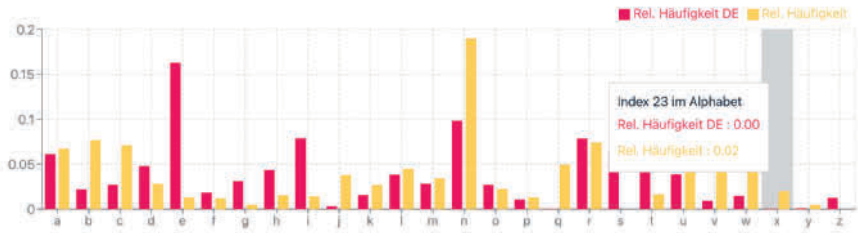
Bei solchen Dechiffrierungsaufgaben ist es hilfreich, gleichzeitig mit dem Lückentext und den bisher erkannten Teilen der Chiffrierungstabelle zu arbeiten. Die bekannten Bezeichnungen von Zeilen und Spalten geben uns wichtige Informationen darüber, welche Zeichen welche Buchstaben kodieren können und welche nicht.

Der Klartext lautet: WIE SCHWIERIG WAR DIESE KNOBELAUFGABE

3. Die Lernumgebung erkunden und die Kompetenzen festigen

BEISPIEL LERNUMGEBUNG «GEHEIMSCHRIFTEN UND DATENSICHERHEIT»

→ klett-online.ch › Caesar-Verschlüsselung I › Monoalphabetische Verschlüsselung



Aufgabe

Der römische Kaiser Gaius Julius Caesar Octavianus hat eine neue Chiffriermethode entwickelt. Voller Stolz zeigt er dir ein Beispiel. Schaffst du es, den Schlüssel zu finden, der für diese Caesar-Verschlüsselung verwendet wurde? Zur Unterstützung haben wir dir oben die Verteilung der relativen Häufigkeiten eingezeichnet.

Hint: Der Schlüssel ist eine Zahl zwischen 0 und 25.

Klartext

EINSYMMETRISCHESKRYPTOSYSTEMISTEINKRYPTOSYSTEMBEIWELC
HEMIMGEGENSATZZUEINEMASYMMETRISCHENKRYPTOSYSTEMBEIDE
TEILNEHMERDENSELBENSCHLUESSELVERWENDENBEIMANCHENSYM
METRISCHENVERFAHRENZB DEASINDDIEBEIDENSCHLUESSELNICHTID
ENTISCHABERKOENNENLEICHTAUSEINANDERBERECHNETWERDENDIE
SERARTIKELBESCHRAENKTSICHAUFDIEDARSTELLUNGVONVERSCHLU

Geheimtext

NRWBHVVNCARBLQNBTAHYCXBHBCNVRBCNRWTAHYCXBHBCNVKNR
FNULQNVRVPNPNWBJCIIDNRWNVJBHVVNCARBLQNWTAHYCXBHBCN
VKNRWNCNRUWNQVNAMNWBNUKNWBLQUDNBBNUENA NWMNWK
NRVJWLQNWBHVVNCARBLQNWENAOJQANWIKRMNJBRWMMRNKNR
MNWBLQUDNBBNUWRLQCRMNWCRBLQJKNATXNWWNWUNRLQCJD
BNRWJWMNAKNANLQWNCFNAVNWMRNBNAJACRTNUKNBLQAJNWT

Alle Lernumgebungen auf **klett-online.ch**